#6

Eur päisches Patentamt

European Patent Office

Office eur péen des brevets

jc678 U.S. PTO 09/504393 02/15/00

Bescheinigung

Certificate

Attestation

Die angehefteten Unterlagen stimmen mit der ursprünglich eingereichten Fassung der auf dem nächsten Blatt bezeichneten europäischen Patentanmeldung überein.

The attached documents are exact copies of the European patent application described on the following page, as originally filed.

Les documents fixés à cette attestation sont conformes à la version initialement déposée de la demande de brevet européen spécifiée à la page suivante.

Patentanmeldung Nr. Patent application No. Demande de brevet n°

99103382.0

Der Präsident des Europäischen Patentamts;
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets
p.o.

Hatten

I.L.C. HATTEN-HECKMAN

DEN HAAG, DEN
THE HAGUE, 13/01/00
LA HAYE, LE

EPA/EPO/OEB Form 1014 - 02.91

THIS PAGE BLANK (USPTO)

Eur päisches Patentamt | European Patent Office | Office européen des brevets

Blatt 2 der Bescheinigung
Sheet 2 of the certificate
Page 2 de l'attestation

Anmeldung Nr.:
Application no.: 99103382.0
Demande n°:

Anmeldetag:
Date of filing: 22/02/99
Date de dépôt:

Anmelder:
Applicant(s):
Demandeur(s):
F.HOFFMANN-LA ROCHE AG
4070 Basel
SWITZERLAND

Bezeichnung der Erfindung:
Title of the invention:
Titre de l'invention:
Beta,beta-carotene 15,15'-dioxygenases, nucleic acid sequences coding therfore and their use

In Anspruch genommene Prioriät(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:
State:
Pays:

Tag:
Date:
Date:

Aktenzeichen:
File no.
Numéro de dépôt:

Internationale Patentklassifikation:
International Patent classification:
Classification internationale des brevets:
/

Am Anmeldetag benannte Vertragstaaten:
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE
Etats contractants désignés lors du depôt:

Bemerkungen:
Remarks:
Remarques:

THIS PAGE BLANK (USPTO)

LEDERER, KELLER & RIEDERER

Patentanwälte - European Patent Attorneys

A. VAN DER WERTH
(1934 - 1974)
DR. FRANZ LEDERER
Dipl.-Chem. München
DR. GÜNTER KELLER
Dipl.-Biol. München
DR. MICHAEL BEST
Dipl.-Chem. München
ANTON FRH. RIEDERER v. PAAR
Dipl.-Ing. Landshut
80538 MÜNCHEN
Prinzregentenstraße 16
Telefon (089) 21 23 99 0
Telefax (089) 21 23 99 22
E-Mail lederer_keller@compuserve.com

22. Februar 1999

EPO - Munich
8
22. Feb. 1999

F. Hoffmann-La Roche AG
124 Grenzacherstraße

CH-4070 Basel

Schweiz

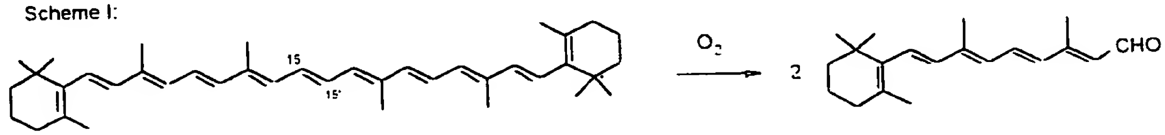
β,β-carotene 15,15'-dioxygenases, nucleic acid sequences coding therefore and their use

The present invention concerns the cloning of β,β-carotene 15,15'-dioxygenase (EC 1.13.11.21), the enzyme responsible for the cleavage of β-carotene leading to vitamin A. The term vitamin A as defined in the present invention comprises a class of compounds including retinal, retinol, 3-dehydroretinol, retinoic acid, the isomers from these compounds as well as retinylesters. Proteins having β,β-carotene 15,15'-dioxygenase activity and nucleic acid sequences coding therefore can be used in different fields including but not limited to diagnostics, the technical production of vitamin A, the generation of transgenic plants in order to produce vitamin A in fruits and vegetables, or gene therapy.

Vitamin A is essential for man and animal and is largely formed in most organisms from its precursor carotenoids which, by themselves, can only be formed in plants, in photo-synthetic active and some other microorganisms. Man and most animals (in particular herbivores and omnivores) are able to convert such

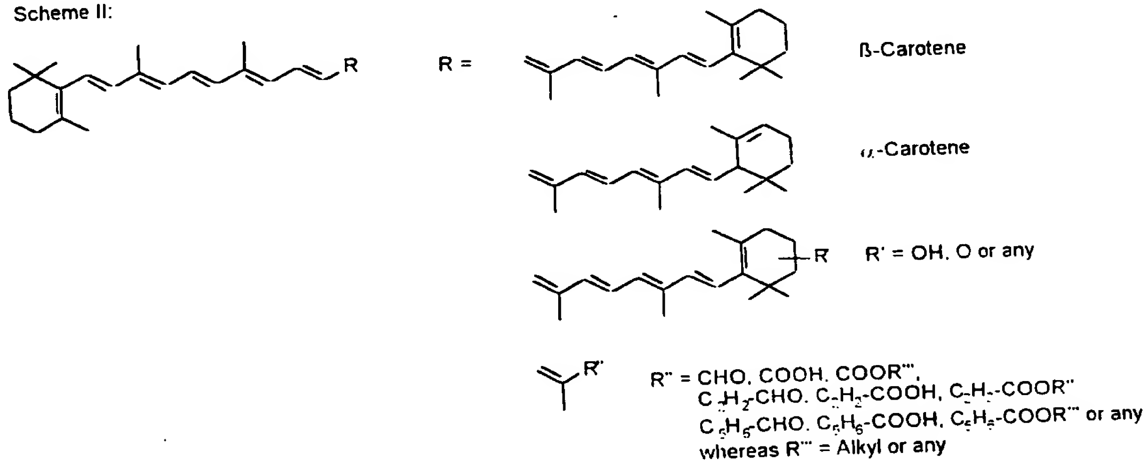
carotenoids, also called provitamins A, enzymatically into vitamin A. The most important enzyme for this process is the β,β-carotene 15,15'-dioxygenase (EC 1.13.11.21). The enzyme is located in the cytosol and forms retinal from β-carotene, as the principal substrate, in presence of oxygen according to scheme 1:

Scheme I:



The enzyme β,β-carotene 15,15'-dioxygenase is characterized by generating 2 mols retinal from 1 mol of β-carotene by central cleavage. But the enzyme is also able to convert a wide range of carotenoids into vitamin A-active compounds, as shown in scheme II:

Scheme II:



Highest known enzymatic activity is found in the intestine of herbivores, especially in duodenum. In other tissues like liver, lung, kidney and brain β,β-carotene 15,15'-dioxygenase is also detectable. Starting in 1955, many attempts have been undertaken to purify and characterize the enzyme by biochemical methods [Goodman (1965 and 1966), Fidge (1969), Laksmanan (1972), Sharma (1977) and Devery & Milborrow (British Journal

of Nutrition (1994) 72, p. 397-414}. However, none of these attempts has been successful. Specific activities of 600 pmol retinal formed/mg protein per hour have not been surpassed.

In the course of the present invention it was possible to purify the chicken enzyme to such a degree that a partial amino acid sequence could be obtained. The enzyme was enriched 226-fold, yielding a specific activity of 2500 pmol/h/mg. On a polyacrylamid gel of fractions from the final gel filtration run 15 bands were visible after Coomassie blue staining. Two bands correlated with the enzymatic activity profile. With the first protein Edman sequencing and with the second MS spectroscopy was performed. Trypic digestion and subsequent MALDI-TOF MS of this latter protein revealed 2 peptides of 11 and 18 amino acids. From this sequence information degenerate PCR primers were designed and synthesized.

With a PCR protocol a 51 bp (base pair) fragment was amplified within the longer peptide. From this sequence a homologous primer was synthesized and used in a second RT-PCR (reverse transcriptase - PCR) to amplify a 597 bp fragment.

This cDNA fragment was radioactively labeled and used for the screening of two positive pools from a chicken expression library in order to isolate the full length cDNA coding for β,β-carotene 15,15'-dioxygenase.

The positive pools were obtained from a cDNA library from chicken duodenum which was screened for β,β-carotene 15,15'-dioxygenase activity in a cellular transactivation assay. By this strategy several positive cDNA pools were identified. By combining the two strategies the gene coding for β,β-carotene 15,15'-dioxygenase could be successfully cloned.

It is an object of the present invention to provide a protein having the vitamin A producing activity of β,β-carotene 15,15'-dioxygenase comprising an amino acid sequence which is identical or homologous to sequence ID no. 1 (shown in Fig. 4) whereby the degree of homology to sequence ID no. 1 is at least 60%.

After having provided the sequence of β,β-carotene 15,15'-dioxygenase isolated from chicken corresponding proteins from different animals like swine, cow, goat, dog, rabbit, poultry, fish and humans can easily be obtained. Since the sequence is known suitable regions of the nucleic acid sequence can be selected as primers for a polymerase chain reaction with a suitable nucleic acid which allows an easy and rapid amplification of the gene coding for the protein.

The present invention comprises therefore not only proteins having an amino acid sequence identical to the sequence given in sequence ID no. 1 but also such proteins which have an amino acid sequence homologous to the sequence ID no. 1. The degree of homology is, however, at least 60%, preferably 70%, more preferably 80% and especially preferred at least 90%. Homology as defined in the present invention means that when the amino acid sequences of two proteins are aligned at least the given percentage is identical. The alignment of the amino acids can be performed with the help of a suitable computer programme which is commercially available, perferably the Wisconsin Sequence Analysis Package GCG (Genetics Computer Group, University Research Park, Madison), Version 9.1, 1997. The remainder of the amino acids may be different. A homology of 90% for example means that 90% of the amino acids of the protein are identical compared with the amino acid sequence given in sequence ID no. 1 whereas 10% of the amino acids may be different. The proteins of the present invention have, however, the biological activity of β,β-carotene

15,15'-dioxygenase which is explained above in more detail. Preferred proteins which are within the scope of the present invention are derived from other mammals or from humans.

Another aspect of the present invention concerns nucleic acid sequences coding for a protein having the biological activity of β,β-carotene 15,15'-dioxygenase. A nucleic acid sequence coding for the enzyme derived from chicken is shown in sequence ID no. 2 (see Figure 3). The nucleic acid sequences of the present invention code for a protein of the present invention or a part thereof. Shorter nucleotide sequences suitable for PCR have a length of at least 20 bases, preferably 25 bases and most preferred at least 30 bases.

The nucleic acid sequences of the present invention can be preferably used as primers for the specific amplification of a gene or part thereof coding for β,β-carotene 15,15'-dioxygenase. Primers can also be used for the specific amplification of 5' nontranslating or 3' nontranslating sequences of the cDNA described above. The nucleic acid sequences of β,β-carotene 15,15'-dioxygenase cDNA can be used as a probe for the detection of the coding as well as for the noncoding regions or parts thereof. The nucleic acid sequences of the present invention can be preferably used as antisense RNA probe for in situ hybridization.

It is especially preferred to use primers and probes having a part of the sequence given in sequence ID no. 2 as primers and/or probes in testkits which can be used for the amplification and/or detection of genes/mRNAs coding for β,β-carotene 15,15'-dioxygenase by polymerase chain reaction (PCR). The selection of suitable parts of the nucleic acid sequence can be performed by the person skilled in the art without difficulties. A nucleic acid sequence used as primer or probe is usually selected from a region which is highly conserved

within the protein. Conserved means that the nucleic acid sequences of such regions of proteins obtained from different species are very similar. On the other hand the preferred nucleic acid sequence should not be present in other nucleic acid sequences which do not code for β,β-carotene 15,15'-dioxygenase, because this might lead to false positive results. By aligning several sequences derived from different species such regions can easily be determined. Although the nucleic acid can be a ribonucleic acid it is more preferred to have desoxyribonucleic acid sequences.

One preferred use in diagnostics is the determination of the β,β-carotene 15,15'-dioxygenase in patients. There is a variability in β-carotene cleavage potential among the population. Humans with low dioxygenase levels (with e.g. mutations or polymorphisms in the gene for β,β-carotene 15,15'-dioxygenase) could be identified and selected for vitamin A supplementation.

A diagnostic kit based on PCR can be designed to detect frequent mutations in the dioxygenase gene. Another diagnostic option is quantification of mRNA by RT-PCR. With this diagnostic tool differences in expression of β,β-carotene 15,15'-dioxygenase in various tissues and in different species can be found.

Since the protein has been expressed and a method for purifying the protein is described in detail in the examples the person skilled in the art can use the protein or peptides derived from the amino acid sequences in order to generate antibodies which do specifically react with the protein. It is either possible to produce polyclonal antibodies by immunizing laboratory animals, like rabbits, sheeps or goats preferably with an adjuvant or monoclonal antibodies by the well-known technique described by Köhler and Milstein [European Journal of

Immunology, 1976, 6 (7), p. 511-519]. The antibodies should specifically react with β,β-carotene 15,15'-dioxygenase in order to avoid an unspecific crossreaction. This means that the antibodies of the present invention should preferably react with an epitope which is present only on a protein of the present invention.

Such antibodies can be preferably used in immunoassays for the detection and/or quantification of β,β-carotene 15,15'-dioxygenase in a test fluid. The test fluid may be a liquid, like serum, obtained from a patient. There are several types of immunoassays which are well-known to the person skilled in the art. Very frequently one antibody, preferably a monoclonal antibody is fixed to a solid phase. This antibody is then brought into contact with the fluid containing the β,β-carotene 15,15'-dioxygenase and after washing it is further reacted with a second monoclonal antibody which binds to another epitope of the enzyme. The second antibody is usually labeled and shows the presence of the sandwich consisting of the antigen and two different antibodies.

The antibodies can also be used in laboratory methods like Western blots or immuno-precipitations. Preferably such antibodies can be used in immunohistochemistry to detect epitopes of β,β-carotene 15,15'-dioxygenase in embedded or fixed tissues or cells of any species of interest.

In a further embodiment of the present invention the β,β-carotene 15,15'-dioxygenase is used for the production of vitamin A whereby the enzyme cleaves enzymatically β-carotene into two molecules of retinal which will subsequently be reduced by retinol dehydrogenase to vitamin A. The β,β-carotene 15,15'-dioxygenase can be used to enzymatically convert β-carotene which may be obtained from plant sources. A preferred

source of β-carotene is the alga Dunaliella bardawil which has a high endogenous level of β-carotene. Suitable algae can be grown conveniently and β-carotene can be purified therefrom at rather low cost. The carotene can be conveniently cleaved enzymatically by using a protein of the present invention. The carotene dioxygenase can preferably be immobilized in order to provide a continuous process.

Another aspect of the present invention concerns the introduction of the gene coding for a protein having β,β-carotene 15,15'-dioxygenase activity into a suitable host cell. The first step in such a method is usually to insert the cDNA into a suitable vector. The vector must fit with the host cell into which the gene should be introduced. There are specific vectors available for bacteria, yeasts, plant cells, insect cells or mammalian cells. Preferably the gene is combined with genetic structures which provide the required genetic regulation like promotors, enhancers, ribosomal binding sites etc.

Systems for the expression of genes encoding carotenoid biosynthetic enzymes in procaryotes, especially in E.coli or Bacillus subtilis or Flavobacter and eucaryotes, e.g. fungi are known in the art and described e.g. in European Patent Application Publication No. (EP) 747 483 or in EP 872 554.

The vector having the gene and the other required genetic structures is then introduced into suitable host cells by well-known methods like transformation, transfection, electroporation or microprojectile bombardment. Depending on the host cell it may be preferred to stably integrate the gene coding for a protein of the present invention into the genome of the host cell. The cells obtained by such methods can then be further propagated and if the cell is a plant cell it is possible to generate therefrom transgenic plants.

In one embodiment of the present invention host cells are plant cells and tomato cells are especially preferred. The technology to produce transgenic tomatoes is well-established and the tomato contains sufficient β-carotene in order to come up with a reasonable vitamin A level after introduction of carotene dioxygenase into the tomato plant. In green pepper, melon or especially carrot the endogenous level of β-carotene is even higher and therefore also these plants are especially preferred.

Another preferred embodiment of the present invention concerns algae. Halotolerant algae may contain high levels of β-carotene. A transfection of such algae with an expression vector comprising the β,β-carotene 15,15'-dioxygenase cDNA leads to a high intracellular vitamin A level which can easily be recovered from such algae by simple purification steps.

In another aspect of the present invention a gene coding for a protein of the present invention can be introduced into mammalian cells and especially into human cells. It is for example possible to insert the gene coding for a β,β-carotene 15,15'-dioxygenase into suitable cells, for example peripheral blood stem cells. Such cells which contain the gene for β,β-carotene 15,15'-dioxygenase may be administered to people having mutations or deletions in the β,β-carotene 15,15'-dioxygenase gene. Such mutations and deletions, respectively, may have the effect that such patients are not able to cleave β-carotene enzymatically. Therefore, such patients always have a low vitamin A level and thus suffer from various developmental and ophthalmological problems. The administration of suitably transfected cells expressing the β,β-carotene 15,15'-dioxygenase to such patients by way of somatic gene therapy is a promising way to improve their situation.

Some of the most important results of the present invention are summarized in the Figures.

Figure 1 shows the result from the last step of purification of β,β-carotene 15,15'-dioxygenase from the small intestine of chicken. The SDS-PAGE pattern and β,β-carotene 15,15'-dioxygenase activity of individual fractions from the gel permeation chromatography run are shown. On the gel, the protein A is marked by an arrow and seemed to correlate best with the β,β-carotene 15,15'-dioxygenase activity. It was therefore chosen for further amino acid sequence analysis. The abbreviations have the following meaning: Std.: molecular weight standard; conc.: concentrate loaded onto the gel permeation chromatography column.

Figure 2 shows schematically the transactivation assay in eukaryotic cells. cDNAs are transfected and expressed in MCF-7 cells. When incubated with β-carotene, a positive pool shows cleavage activity. The cleavage product retinal is further oxydized to retinoic acid (RA) which binds to the endogenous receptor. The receptor/ligand complex binds to the response element on the reporter plasmid and leads to an enhanced transcription of the luciferase gene. The luminescene signals are detected in a luciferase assay with a sensitive CCD camera.

Figure 3 shows the cDNA sequence (sequence ID no. 2) for β,β-carotene 15,15'-dioxygenase which has a length of 3090 base pairs excluding the poly A tail. 132 base pairs are 5' nontranslating sequence, the coding sequence has 1578 base pairs and the 3' nontranslating sequence 1380 base pairs, respectively. A putative poly A signal is found at position 3073.

Figure 4 shows the derived amino acid sequence (sequence ID No. 1) of β,β-carotene 15,15'-dioxygenase derived from chicken having 526 residues. The amino acid sequence is given in the one letter code.

Figure 5 shows a comparison of the β,β-carotene 15,15'-dioxygenase amino acid sequence with a protein having the designation RPE65 which was found by a sequence comparison in EMBL Genbank as the protein having the highest homology to the β,β-carotene 15,15'-dioxygenase of the present invention.

The present invention is further illustrated in the following examples:

Example 1:

Assay of β,β-carotene 15,15'-dioxygenase activity

For the tests the following solutions were prepared:

a) Solution 1 (mixed micelle solution): Glycocholic acid (1.16 g) was dissolved in 5 ml H_2O under stirring and by drop-wise addition of 5 N NaOH. After pH was adjusted to 6.8-7.2 with acetic acid and the volume increased to 10 ml with H_2O, 80 mg of asolectin (Fluka) were added and dissolved under stirring.

b) Solution 2 (substrate solution): 500 µl of an α-tocopherol solution (10 mg/ml in hexane) and 235 µg of a β-carotene solution (80 µg/ml of pure all-E-β-carotene in

benzene) were mixed in a glass vial, protected from light and the solvents evaporated under a gentle nitrogen stream. 1 ml of solution 1 was added under vortexing and eventually a few short ultrasonic bursts until a clear solution occurred.

c) Solution 3 (homogenization buffer): 100 mM KH_2PO_4 adjusted with 5 N KOH to pH 7.8 and containing 4 mM $MgCl_2$, 6 H_2O and 30 mM nicotinamide.

d) Solution 4 (GSH solution): 60 mg/ml reduced glutathione dissolved in solution 3.

e) Solution 5 (standard solution): 10 µg/ml vitamin A acetate in hexane/chloroform 9:1.

Activity assay: 2 ml enzyme preparation (approx. 4 mg protein, assayed by BCA protein assay, Pierce Chemicals) were placed in a light-protected glass vial in a shaking water bath (30 min., 37 °C). 0.2 ml solution 4 was added and the reaction was started after 2 min. of temperature equilibration by addition of 50 µl solution 2. After 3 h, the reaction was stopped by placing the vials on ice and subsequent addition of 1 ml acetonitrile followed by 5 ml chloroform. The vials were vortexed 3 times for 7 s and phase separation was enforced by centrifugation for 5 min. at 5000 g. Extraction was repeated twice with 0.6 ml chloroform. The combined chloroform phases were

evaporated and resolubilized in 200 µl solution 5 under short sonication. Insoluble material was removed by filtration through 0.45 µm filters. An aliquot of 20 µl was separated by HPLC on a reversed phase C_{18} column (Lichrospher 100, 5 µm, 12.5 cm x 4.6 mm, Bischoff Chromatography, Leonberg, Germany; 1 ml/min, column temperature 25°C) with a discontinuous, optimized gradient of acetonitrile/tetrahydrofuran/(1% ammonium acetate in H_2O) from 50:20:30 (eluent A) to 50:44:6 (eluent B). These conditions allow complete separation of β,β-carotene and retinal as well as apo-β-carotenals and retinoic acids. Calibration curves were made for both β,β-carotene and retinal in the concentration ranges 2-40 and 1-10 ng/µl, respectively, and were correlated to the value of vitamin A acetate which served as an internal standard. Enzymatic activity was expressed as the amount of retinal liberated in the activity assay during 3 h of incubation at 37 °C (100% = 17.6 nmol).

Example 2:

Purification of β,β-carotene 15,15'-dioxygenase

Purification was done as rapidly as possible, and all buffers and equipments were cooled to 4 °C.

Solution 6 (protease inhibitor-containing homogenisation buffer): 125 mM benzamidine•HCl, 250 mM 6-aminocaproic acid and 125 µM soybean trypsin inhibitor were dissolved in H_2O by sonication. A 4 ml-aliquot of this solution was mixed with 100 ml of solution 3.

Solution 7: 10 mM KH_2PO_4, 1 mM reduced glutathione, pH 7.8.

Solution 8 (eluent A, phenyl-Sepharose chromatography): 10 mM KH_2PO_4, 1 mM reduced glutathione, 0.5 M $(NH_4)_2SO_4$, pH 7.8.

Solution 9 (eluent B, phenyl-Sepharose chromatography): 10 mM KH_2PO_4, 1 mM reduced glutathione, 10% glycerol, pH 7.8.

Solution 10 (eluent B, Poros HQ chromatography): 10 mM KH_2PO_4, 1 mM reduced glutathione, 0.5 M NaCl, 10% glycerol, pH 7.8.

Solution 11 (elution buffer for gel permeation chromatography): 50 mM KH_2PO_4, 1 mM reduced glutathione, 150 mM NaCl, 10% glycerol, pH 7.8.

Laying hens at age of 20-24 weeks (strain Lohmann LSL, Hatchery Wuethrich, CH-3123 Belp, Switzerland) were kept on a pigment-free chicken diet (Kliba 3179, Kliba, CH-4303 Kaiseraugst, Switzerland). The animals were killed by decapitation and the first 20 cm of the duodenal loop was removed, separated from

pancreas and rinsed with 40 mL each of 0.9% NaCl solution. The intestines were immediately frozen in dry ice and stored at -80°C until use.

Ten intestines were thawed on ice in approx. 2 h and opened length-wise in an ice cooled Petri dish. The mucosa was scraped off with a slide, weighed and homogenized in a Teflon-glass Potter-Elvehjem homogeniser in 4 volumes of solution 6 with six strokes. Upon centrifugation at 62000 rpm for 1 h, the clear supernatant was divided into 32 aliquots of 15 mL each. From these preparations an ammonium sulphate fractionation was made. The precipitate obtained from the 20-45% step was centrifugated at 5000 g for 10 min. and the pellet was stored at -80°C for further use.

Ten aliquots of the $(NH_4)_2SO_4$ pellet were dissolved in 150 ml of solution 7, sterile-filtered and loaded on a HiLoad 26/10 phenyl-Sepharose High Performance column (column volume 53 ml; Pharmacia, Uppsala, Sweden) equilibrated with solution 8. Proteins were eluted at a flow rate of 8 ml/min with a steep gradient over 1 column volume (CV) from solution 8 to solution 9. β,β-Carotene 15,15'-dioxygenase eluted at a conductivity of < 15 mS/cm, but only fractions with a conductivity of < 1 mS/cm were pooled and directly loaded onto a 30 ml Blue Sepharose 6 Fast Flow column (Pharmacia) equilibrated with solution 9. β,β-Carotene 15,15'-dioxygenase activity eluted (at a flow rate of 8 ml/min) in the

break-through fractions which were (again) directly loaded onto a 20 ml Poros HQ/M anion exchange chromatography column (PerSeptive Biosystems, Framingham, MA, USA) equilibrated with solution 9. β,β-Carotene 15,15'-dioxygenase was eluted at a flow rate of 15 ml/min with a linear gradient over 18 CV from solution 9 to solution 10. Activity was detected in the gradient in a conductivity range of 10-20 mS/cm. The pooled fractions (70 ml) were concentrated to ~1.3 ml with Ultrafree-15 filter units (MW cut-off 50,000; Millipore, Bedford, MA, USA). An aliquot of the concentrate (500 µl) was loaded onto a Superdex 200 HR 10/30 gel filtration column (CV 24 ml; Pharmacia) and eluted at a flow rate of 0.5 ml/min with solution 11. Aliquots of each fraction were used for activity assays (see example 1) and, upon concentration, for SDS-PAGE (with MOPS running buffer) on 10% NuPAGE gels (Novex, San Diego, CA, USA). The results of this experiment are shown in Fig. 1 and Table 1.

Purification step	Total protein [mg]	Total activity [nmol/h]	Yield [%]	Specific activity [pmol/(h·mg)]	Purification factor
$(NH_4)_2SO_4$ pellet	779	8.61	100	11.0	–
phenyl-Sepharose	80.8	8.27	96.1	102	9.27
Blue Sepharose	16.0	8.86	103	554	50.1
Poros HQ	1.56	3.90	45.3	2500	226

Table 1: Summary table for the purification of β,β-carotene 15,15'-dioxygenase starting with 10 aliquots of the $(NH_4)_2SO_4$ pellet (means of 3-4 measurements).

Example 3:

Amino acid sequence information for β,β-carotene 15,15'-dioxygenase

For amino acid sequence analysis, fractions of the gel filtration run (as shown in Fig. 1) were separated on a 8-16% Tris/glycine gel (Novex), and the proteins transferred to an Immobilon P^{SQ} membrane (Millipore) and stained with amido black.

Since protein A proved to be N-terminally blocked, multiple aliquots of fraction 18 from the gel filtration run (see Fig. 1) were separated on a 10% Tris/glycine gel (Novex), and the gel was stained with Colloidal Coomassie Blue (Novex). The band corresponding to protein A was excised from the gel, and the protein digested in-gel with trypsin. The tryptic digest was separated by micro-bore reversed-phase HPLC on a 150 x 1.0 mm Vydac C_{18} column (Vydac, Hesperia, CA, USA). Peptides were eluted with an acetonitrile gradient in 0.1% trifluoroacetic acid, and peptide containing fractions were collected for further analysis. Two fractions could be identified by MALDI-TOF-MS (Voyager Elite, PerSeptive Biosystems) to contain one single peptide each. N-terminal Edman degradation revealed the following sequences:

(1) Ala-Glu-Val-Gln-Gly-Gln-Leu-Pro (Seq. ID No. 3)

(2) Asn-Lys-Glu-Glu-His-Pro-Glu-Pro-Ile-Lys-Ala-Glu-Val-Gln-Gly-Gln-Leu-Pro (Seq. ID No. 6)

Note that the last 8 amino acids of peptide (2) correspond to peptide (1).

Example 4:

Cloning the full length cDNA for the β,β-carotene 15,15'-dioxygenase

A) RNA isolation:

4 week old Vedette chicken were killed, the duodenum was removed, washed with sterile PBS and cut open with scissors. The mucosal layer was scraped off with a glass slide, weighed and homogenized immediately with a Polytron in 1 ml of Trizole reagent (Life Technologies) per 100 mg of tissue. Then the standard protocol from Life Technologies was followed. Poly A mRNA was isolated by the polyATtract mRNA Isolation kit from Promega Corporation, Madison.

B) PCR and RT-PCR:

In the peptide sequence NKEEHPEPIKAEVQGQLP (peptide 2 of Example 3) (Seq. ID No. 7) two degenerate primers were designed: In order to have a lower degeneracy the base Inosin was used in one and in two wobble positions, respectively.

5' primer: 5'AAC AAR GAR GAS CAY CCI GA 3' (Seq. ID No. 8) (20 mer with a degeneracy of 16x)

3′ primer: 5′SAG CTG ICC CTG IAC YTC SGC 3′ (Seq. ID No. 9) (21 mer with a degeneracy of 8x)

R= A or G, S= C or G, Y= C or T

The oligos were synthesized on a Pharmacia Gene Assembler Plus using standard phosphoramidite chemistry. Deprotection was done with 1 ml conc. ammonium hydroxide solution (Applied Biosystems) and final desalting was performed with a NAP 10 column (Amersham Pharmacia Biotech).

For PCR 100 ng chicken duodenal cDNA were taken as template and the following steps performed: 94°C 30′′; 52°C 30′′; 72°C 1′ for 40 cycles. The resulting band of 51 bp was cut out from a 10% polyacrylamide gel, electroeluted on DEAE paper at 300 V for 1.5 h, eluted from the DEAE paper once with 40 µl and twice with 30 µl 1.5 M NaCl, 5 mM Tris, 0.5 mM EDTA, precipitated with 2.5 volumes of ethanol 100% and 1 µg glycogen, washed with 0.5 ml of 80% ethanol, dried and dissolved in 20 µl TE (10 mM Tris, 1 mM EDTA).

The resulting fragment of 51 bp was cloned into pGEM-T Easy, a commercially available T/A cloning vector (Promega Corporation, Madison). The corresponding cDNA sequence was determined by automated fluorescent sequencing on a Vistra DNA Sequencer 725 (Amersham Pharmacia Biotech).

From the above DNA sequence a homologous forward primer was derived:

5′ TCT<u>GAATTC</u>CGGAGCCCATAAAAGC 3′ (primer dioxy12) [Seq. ID No. 10]

At the 5′ end an EcoRI site (underlined sequence) was introduced; the following 17 nucleotides are homologous to the previously obtained dioxygenase sequence.

In a RT-PCR reaction a polyT/Not primer (commercially available from Invitrogen, San Diego) was used as reverse primer together with primer dioxy12.

One tube RT-PCR kit from Boehringer Mannheim was taken and the corresponding protocol followed:

mix 1:

18.3 µl	H_2O
2.5 µl	DTT (100 mM)
1.0 µl	dNTPs (10 mM)
1.0 µl	oligo dT/Not (0.2 µg/µl) (3' primer)
1.0 µl	dioxy12 (5' primer) (20 µM)
0.2 µl	RNAse inhibitor (40 U/µl)
1.0 µl	chicken duodenal total RNA (2.2 µg/µl)
25.0 µl	

mix 2:

14.0 µl	H_2O
10.0 µl	RT-PCR buffer 5x
1.0 µl	enzyme mix (AMV RT, Taq and Pwo DNA Polymerase)
25.0 µl	

The 2 mixes were combined and the PCR protocol started on a MJ Research PTC200 DNA Engine.

50°C	30'	
94°C	2'	
94°C	30''	
57°C	30''	10 cycles
68°C	45''	
94°C	30''	
62°C	30''	25 cycles
68°C	45'' + 3'' /cycle	
68°C	7'	
4°C	over night	

With this RT-PCR protocol a band of 597 bp was amplified from chicken total duodenal RNA. The PCR band was isolated from a 1% agarose gel, cloned into pGEM-T Easy cloning vector and

subsequently sequenced. The original peptide is present in the sequence as well as an open reading frame over the whole sequence of 597 bp.

C) Chicken cDNA-library:

From chicken duodenal polyA$^+$ RNA cDNA was made with the Copy Kit (Invitrogen, San Diego) using a modified Gubler-Hoffman procedure. The cDNA was size-selected (0.9-5.5 kb) and cloned into the eukaryotic expression vector pcDNA1.1/Amp (Invitrogen).

Electroporation into E.coli Top10 was done with a Bio-Rad Gene Pulser II system following the standard protocol. This resulted in a cDNA library of 480'000 individual clones. The library was splitted into 250 pools with 1500-2500 individual clones each. Each pool was amplified in 100 ml LB medium: Bacterial growth was stopped at OD 0.8-1.0 by adding chloramphenicol to a final concentration of 170 µg/ml. Incubation was continued over night in order to increase the DNA amount.

D) Activity screening of the chicken cDNA library:

90 of the above pools were tested for activity in a transactivation assay based on the detection of retinoic acid which is produced in eukaryotic cells after β-carotene cleavage. The principle of the activity test is shown in Fig. 2.

5 µg of DNA from each pool were transfected with 20 µg of lipofectin (Life Technologies) into a reporter cell line bearing a luciferase reporter plasmid with a RARE (retinoic acid response element) in front of the tk promoter (Herpes simplex thymidine kinase promoter). Transfections were done for 7 h under serum free conditions. After 7 h the transfection mix was removed and RPMI medium with 10% charcoal treated FCS (fetal calf serum) was added. After 20 h of incubation β-carotene (β-carotene 10% CWS, F. Hoffmann-La Roche Ltd.) or a placebo formulation were added to

the culture medium to a final β-Carotene concentration of 5 μM. Incubation was continued for 18 h. Then cells were washed with PBS, and luciferase expression was measured after substrate addition with a nitrogen cooled slow scan CCD camera (AstroCam Ltd.) Exposure time usually was 8 min. Analysis was done with the Image Pro Plus 3.0 software package (Media Cybernatic, Maryland). 3 pools were strongly positive, 7 pools showed weaker, but detectable activity.

One of the positive pools was plated on a square agar plate. 2 filters (nylon membranes, Gene Screen, NEN Research Products, Boston) were processed and screened with the radioactively ([α^{32}P]dATP, Amersham) labeled 597 bp PCR-fragment. From 9500 colonies screened, 14 were double positive. From 36 colonies picked, 5 showed the same pattern after restriction site analysis. 2 clones were sequenced from the 5′ end and the original 51 bp sequence was found. Subsequently the whole cDNA was sequenced and confirmed twice.

All molecular biological precedures were done according to Sambrook, Fritsch and Maniatis, Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, if not otherwise mentioned.

The obtained cDNA sequence is shown in Figure 3 and the amino acid sequence deduced therefrom in Figure 4.

Figure 4 shows the derived amino acid sequence having 526 residues.

Example 5: Sequence comparison:

By sequence comparison with the EMBL Genbank a high homology between the known protein RPE65 (Hamel et al., J.Biol.Chem. [1993] p. 15751-15757) and the β,β-carotene 15,15′ dioxygenase was found. A homology of 55.5% on the aminoacid level was found. The sequence alignment is shown in Figure 5.

F. Hoffmann-La Roche AG

Claims

1) Protein having β,β-carotene 15,15'-dioxygenase activity comprising an amino acid sequence which is identical or homologous to sequence ID no. 1 whereby the degree of homology to sequence ID no. 1 is at least 60%.

2) Protein according to claim 1 wherein the degree of homology is at least 70%.

3) Protein according to claim 1 wherein the degree of homology is at least 80%.

4) Protein according to claim 1 wherein the degree of homology is at least 90%.

5) Protein according to any one of claims 1 to 4 wherein the amino acid sequence is derived from chicken.

6) Nucleic acid sequence coding for a protein according to any one of claims 1 to 5.

7) Nucleic acid sequence according to claim 6 which comprises at least a part of the sequence of sequence ID no. 2.

8) Nucleic acid sequence according to claim 7 wherein the oligonucleotide has a length of at least 20 bases.

9) Nucleic acid sequence according to any one of claims 6 to 8 wherein the nucleic acid is desoxyribonucleic acid.

10) Nucleic acid sequence according to any one of claims 6 to 8 wherein the nucleic acid is an antisense ribonucleic acid.

11) Primer for the specific amplification of a gene coding for a protein having β,β-carotene 15,15'-dioxygenase activity which comprises a nucleic acid sequence according to any one of claims 6 to 9.

12) Probe for the detection of a gene coding for a protein having β,β-carotene 15,15'-dioxygenase activity which comprises a nucleic acid sequence according to any one of claims 6 to 9.

13) Testkit for the amplification and/or detection of a gene or a part thereof coding for β,β-carotene 15,15'-dioxygenase comprising at least one primer according to claim 11 and/or at least one probe according to claim 12.

14) Antibody characterized in that said antibody specifically reacts with a protein according to any one of claims 1 to 5.

15) Immunoassay for the detection and/or quantification of β,β-carotene 15,15'-dioxygenase which comprises at least one antibody according to claim 14.

16) Process for the production of vitamin A wherein β-carotene is enzymatically cleaved by a protein according to any one of claims 1 to 5.

17) Method for the introduction of the β,β-carotene 15,15'-dioxygenase cDNA into a host cell wherein a cDNA coding for a protein according to any one of claims 1 to 5 is inserted into a vector suitable for the host cell and said vector is introduced into the host cell.

18) Method according to claim 17 wherein the host cell is a plant cell.

19) Method according to claim 17 wherein the host cell is a prokaryotic cell.

20) Method according to claim 17 wherein the host cell is a yeast cell or another fungus.

21) Method according to claim 17 wherein the host cell is an alga.

22) Method according to claim 17 wherein the host cell is a mammalian cell.

23) Method according to claim 22 wherein the mammalian cell is a human cell.

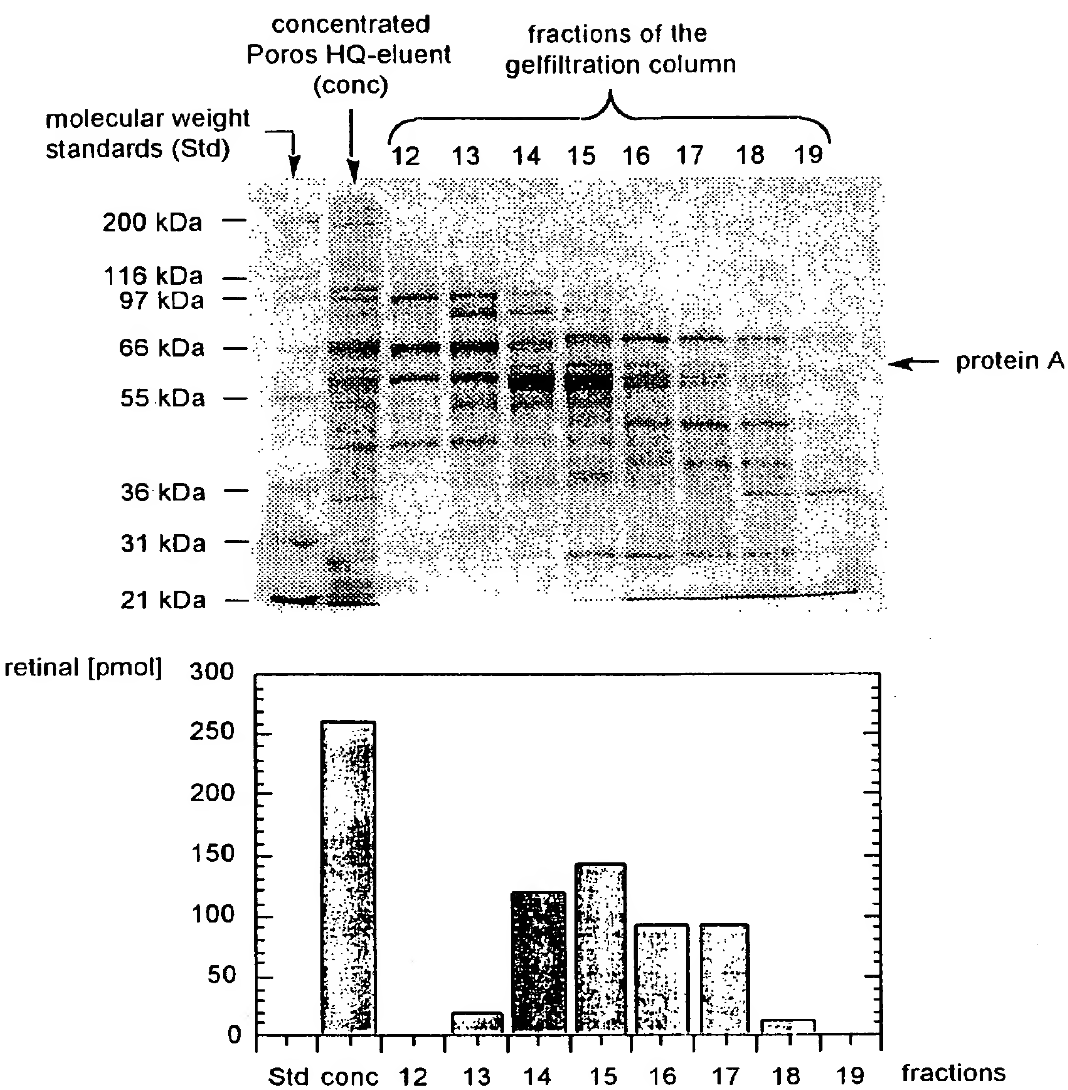
24) Host cell obtainable by a method according to any one of claims 17 to 23.

25) Host cell according to claim 24 which comprises a β,β-carotene 15,15′-dioxygenase cDNA obtained from another species.

THIS PAGE BLANK (USPTO)

F. Hoffmann-La Roche AG

Figure 1



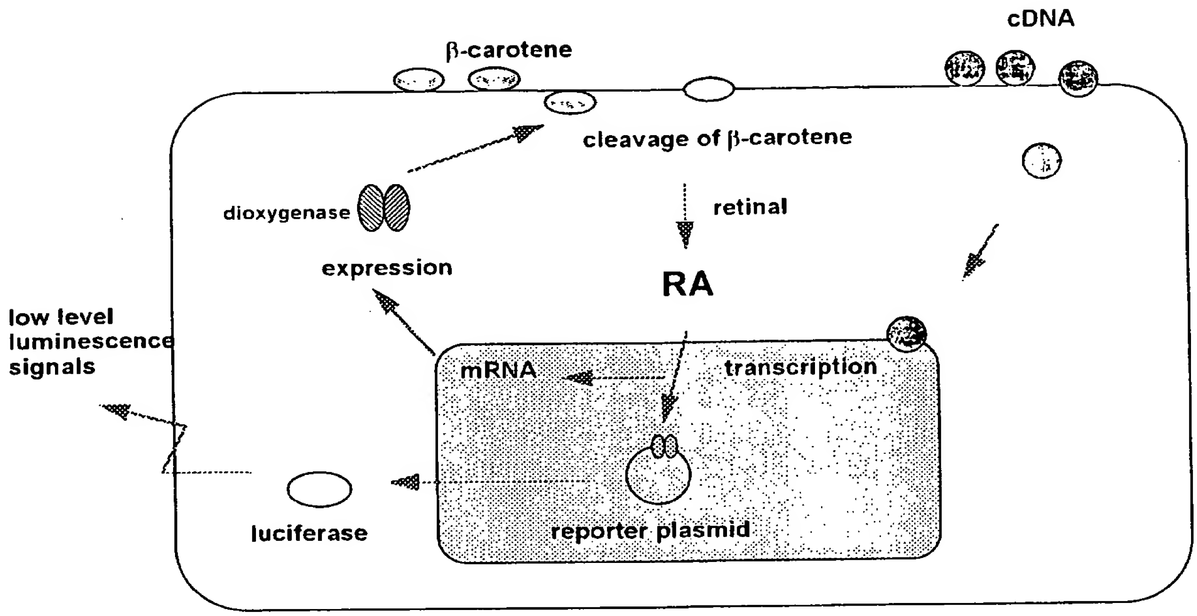

cDNA
β-carotene
cleavage of β-carotene
dioxygenase
retinal
expression
RA
low level luminescence signals
mRNA
transcription
luciferase
reporter plasmid

Figure 2

Figure 3 Seq. ID No. 2

```
   1 CGGATCCACT AGTAACGGCC GCCAGTGTGG TGGAATCCAT CCTTCTATGT
  51 AACAGGAAAG AGCTGTTCTT AGCCCAGAGA GGAGGGCACC GTACGCCTGC
 101 AGGAGCAGCT GGGTAGAGGA CACAGGAGAG CGATGGAGAC AATATTTAAC
 151 AGAAACAAAG AAGAGCATCC AGAGCCCATA AAAGCTGAGG TGCAAGGTCA
 201 GTTGCCCACT TGGTTGCAAG GGGTACTTCT CCGAAATGGC CCAGGGATGC
 251 ACACAATAGG GGACACTAAA TACAACCACT GGTTTGATGG CTTGGCTCTG
 301 CTGCACAGCT TCACGTTTAA AAATGGTGAA GTTTACTACA GAAGTAAGTA
 351 CCTCCGAAGT GACACATACA ACTGCAATAT AGAAGCAAAC CGAATCGTGG
 401 TGTCTGAGTT TGGAACCATG GCTTATCCGG ATCCATGCAA AAACATATTT
 451 GCCAAGGCAT TCTCATACTT ATCTCACACC ATTCCTGAGT TCACGGACAA
 501 CTGCCTGATC AACATTATGA AAACTGGGGA TGATTATTAT GCTACCAGTG
 551 AGACTAACTT CATCAGAAAA ATTGATCCAC AGACTCTGGA GACACTAGAT
 601 AAGGTAGACT ACAGCAAATA TGTAGCTGTA AACTTGGCAA CTTCTCACCC
 651 ACACTATGAC AGTGCTGGAA ATATTCTCAA CATGGGTACT TCAATTGTTG
 701 ATAAAGGGAG AACAAAATAT GTTCTCTTTA AGATCCCTTC CTCTGTACCA
 751 GAAAAAGAAA AGAAGAAATC TTGTTTTAAA CACCTGGAAG TAGTATGCTC
 801 CATCCCTTCT CGCTCCCTGC TCCAACCAAG CTACTACCAC AGCTTTGGAA
 851 TCACAGAAAA TTATATTGTG TTCATAGAGC AGCCATTTAA ACTGGATATT
 901 GTCAAACTGG CAACTGCCTA CATCCGAGGT GTGAACTGGG CTTCCTGCCT
 951 TTCCTTTCAT AAGGAGGATA AGACGTGGTT TCACTTTGTA GACAGAAAGA
1001 CGAAAAAAGA AGTATCCACC AAGTTTTACA CTGATGCTTT GGTGCTTTAT
1051 CACCACATAA ATGCTTACGA AGAAGATGGC CACGTTGTTT TTGATATCGT
1101 TGCCTACAGA GACAATAGCT TGTACGATAT GTTTTACTTA AAAAAACTGG
1151 ACAAAGACTT TGAAGTGAAC AACAAGCTTA CCTCCATCCC AACCTGCAAG
1201 CGCTTTGTTG TGCCTCTGCA GTATGACAAG GATGCAGAAG TAGGTTCTAA
1251 TTTAGTCAAA CTTCCAACTT CCGCAACTGC TGTAAAAGAA AAAGATGGCA
```

1301 GCATCTATTG TCAACCTGAA ATATTATGTG AAGGGATAGA ACTGCCTCGT

1351 GTCAACTATG ACTACAATGG CAAAAAATAC AAGTATGTCT ATGCAACAGA

1401 AGTCCAGTGG AGCCCAGTTC CTACAAAGAT TGCAAAACTG AATGTCCAAA

1451 CAAAGGAAGT ACTGCACTGG GGAGAAGACC ACTGCTGGCC CTCAGAGCCC

1501 ATCTTTGTTC CCAGCCCCGA TGCAAGAGAA GAGGATGAAG GTGTTGTTTT

1551 GACCTGTGTT GTGGTGTCTG AGCCAAATAA AGCACCCTTC CTACTCATCT

1601 TGGATGCTAA AACATTCAAA GAATTGGGCC GAGCCACAGT TAACGTAGAA

1651 ATGCATCTGG ACCTGCATGG GATGTTTATA CCACAGAATG ATTTGGGGGC

1701 TGAGACGGAA TAAAACGCTA TTGATCCGAC TACACAAACT GAGACAACTT

1751 TCTACTGAAC ATGAGTTAAT ATCCCTTTTA CCATTCAAGA ACAACCATAT

1801 AACGACACAA AATGACTATG TATAATCTCT TAAATAATAG ATATAATCCT

1851 TTTAAGGCAC AGCGATGAGT TTTACTACAG GTAACGATAT GCACAACTGG

1901 CATATAACTA TTCCAAAAGA AGAAGAACGA TCAGTGTTTT AGAAGTGCTA

1951 ATGTTGTACA TAACGGCGGC AGAGGGAACA GGAGAGAAAG GTAACGGGAA

2001 TATTTAATAG AATATAGATT TCTGAGCAAA TGAAGTGCAG TATTTATGGT

2051 GTGATGCATG GCATGAGTCA CATAGGTCTG CAGCTCATGT ATCTTTTAGA

2101 GATCGTTTCA AGATTGCAGC TTGTGATGCA AGTTTTCTCC AGCCAGAAAA

2151 CCTCATTTTA AACCATCTGC TACTGGTAAT TCATACCAAT GCATTTTCTT

2201 GGTGCTCGAT TTACACTATA ACCAAAGTTA AGTATTACAT TCAGGTGCTA

2251 CAACTTTCTA ATTTACAACC GAAACAAACA AGCAAACAGC ACTTGCTTTG

2301 CTAATAACCC CATGGTGTAT TTTTCCTTTT TATGATGACA AAACCAAGTA

2351 CATATGGTTT TATGTAGCAT TCAATTATAC TTCAGTGCTA TTCCATCCTA

2401 ATGTTATAAG CAATTTGTAT TTAAATCAGT TTTCCTTGAG AATATCTGAC

2451 ATAACATTTT GTGTAATGAG ATGACTATGT TGTCTAAAGA TGAACAGGAA

2501 TGTATCTTTT ATTAGTATTG TTAATTGTGT TACTAATACT ATGCATATGA

2551 ATGAGAGCAA TGTATTTCTA GGAGAACTCA GATATACATT CAACAATTTC

2601 TGTAGGTGAA AATGCATTTA CTGATGAAAG TTGAATCGTT AATGAGGGAG

2651 AAAACTGGGT ATCCATCCAT CCAACTATGT TAGGTGTTCA CCTGGTCTGT

2701 ATGTGACACC ACGCTGTTTG GGTATCTCTC ACTTTCACAT ACCTGTTCTC

2751 ATGGTTTCTG CTACTCACTG TATTTTGCAG GAGAGAAACA AAATGAAATC

2801 ACTGTCACTT ACTATCGCCC CATCACATAA GAACAATGGG GCTTTGGTGA

2851 CTTGTTCATG ATTACATAAG ATGTTTGCAG CAGAGCAGCA ATAGAACCAA

2901 CACCATCCAC AGTTCTTGCT TGCTCTGTTA TGACTCCCTT TGCTGTCTTT

2951 ATGGTTTGCA TGTATGAAGA ATACACTGCC TAATTCTAAT GTTAAAAAGT

3001 CACTGGGGTC AGATCTAGAG CTTAAGTAAG CAGTCTGGGG TTTTCAAATG

3051 TTTATATGTT CCATAAAATG GAAATAAACA CCTCCATAAT AAAAAAAAAA

3101 AAAAAAAAAA A

Figure 4 **Seq. ID No. 1**

1	METIFNRNKE	EHPEPIKAEV	QGQLPTWLQG	VLLRNGPGMH	TIGDTKYNHW
51	FDGLALLHSF	TFKNGEVYYR	SKYLRSDTYN	CNIEANRIVV	SEFGTMAYPD
101	PCKNIFAKAF	SYLSHTIPEF	TDNCLINIMK	TGDDYYATSE	TNFIRKIDPQ
151	TLETLDKVDY	SKYVAVNLAT	SHPHYDSAGN	ILNMGTSIVD	KGRTKYVLFK
201	IPSSVPEKEK	KKSCFKHLEV	VCSIPSRSLL	QPSYYHSFGI	TENYIVFIEQ
251	PFKLDIVKLA	TAYIRGVNWA	SCLSFHKEDK	TWFHFVDRKT	KKEVSTKFYT
301	DALVLYHHIN	AYEEDGHVVF	DIVAYRDNSL	YDMFYLKKLD	KDFEVNNKLT
351	SIPTCKRFVV	PLQYDKDAEV	GSNLVKLPTS	ATAVKEKDGS	IYCQPEILCE
401	GIELPRVNYD	YNGKKYKYVY	ATEVQWSPVP	TKIAKLNVQT	KEVLHWGEDH
451	CWPSEPIFVP	SPDAREEDEG	VVLTCVVVSE	PNKAPFLLIL	DAKTFKELGR
501	ATVNVEMHLD	LHGMFIPQND	LGAETE		

Figure 5 Seq ID No. 4 and Seq ID No. 5

```
10  EEHPEPIKAEVQGQLPTWLQGVLLR..NGPGMHTIGDTKYNHWFDGLALL 57
20  EELSSPLTAHVTGRIPLWLTGSLLRCFTGPGLFEVGSEPFYHLFDGQALL 69

58  HSFTFKNGEVYYRSKYLRSDTYNCNIEANRIVVSEFG..TMAYPDPCKNI 105
70  HKFDFKEGHVTYHRRFIRTDAYVRAMTEKRIVITEFGFTTCAFPDPCKNI 119

106 FAKAFSYLSHTIPEFTDNCLINIMKTGDDYYATSETNFIRKIDPQTLETL 155
120 FSRFFSYFRGV..EVTDNALVNVYPVGEDYYACTETNFITKINPETLETI 167

156 ..DKVDYSKYVAVNLATSHPHYDSAGNILNMGTSIVDKGRTKYVLFKIPS 203
168 FTKQVDLCNYVSVNGATAHPHIENDGTVYNIGNCFGKNFSIAYNIVKIPP 217

204 SVPEKEKKKSCFKHLEVVCSIPSRSLLQPSYYHSFGITENYIVFIEQPFK 253
218 LQADKEDPISKFTS.EIVVQFPCSDRFKPSYVHSFGLTPNYIVFVETPVK 266

254 LDIVKLATAY.IRGVNWASCL.SFHKEDK.TWFHFVDRKTKKEVSTKFYT 300
267 INLFKFLSSWSLWGANYMDCFESFTNETMGVWLHIADKKRKKYLNNKYRT 316

301 DALVLYHHINAYEEDGHVVFDIVAYRDNSL...YDMFYLKKLDKDFE... 344
317 SPFNLFHHINTYEDNGFLIVDLCCWKGFEFVYNYFTLYLANLRENWEEVK 366

345 VNNKLTSIPTCKRFVVPLQYDKDAEVGSNLVKLP.TSATAV..KEKDGSI 391
367 KNARKAPQPEVRRYVLPLNIDK.ADTGKNLVTLPNTTATAILCSDEFTTI 415

392 YCQPEILCEG....IELPRVNYD.YNGKKYKYVYATEVQWSPVPTKIAKL 436
416 WLEPEVLFSGPRQAFEFPQINYQKYCGKPYTYAYGLGLNHF.VPDRLCKL 464

437 NVQTKEVLH..WGEDHCWPSEPIFVPSPDAREEDEGVVLTCVVVSEPNKA 484
465 NVKTKETWFTVWQEPDSYPSEPIFVSHPDALEEDDGVVLSVVVSPGAGQK 514

485 P.FLLILDAKTFKELGRA..TVNVEMHLDLHGMF 515
515 PAYLLILNAKDLSEVARAEFTVEINIPVTFHGLF 548
```

THIS PAGE BLANK (USPTO)

EPO - Munich
8
22. Feb. 1999

F. Hoffmann-La Roche AG February 22, 1999

Abstract

Nucleic acid sequences coding for a protein having β,β-carotene 15,15'-dioxygenase activity and their uses in diagnosis, the synthesis of vitamin A and methods for the introduction of the β,β-carotene 15,15'-dioxygenase cDNA into host cells are disclosed.

THIS PAGE BLANK (USPTO)